2175
#13
KWS
2-3-03

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

First Named Inventor	: Simon H. Corston-Oliver et al.	Group Art Unit: 2175
Appln. No.	: 09/336,200	Examiner: T. N. Pardo
Filed	: June 18, 1999	
For	: SYSTEM FOR IMPROVING THE PERFORMANCE OF INFORAMTION RETRIEVAL-TYPE TASKS BY INDENTIFYING THE RELATIONS OF CONSTITUENTS	
Docket No.	: M61.12-0099	

RESPONSE

Commissioner for Patents
Washington, D.C. 20231

Sir:

This is in response to the Office Action mailed October 25, 2002. In the Office action claims 1-35, 41, 43 and 62-66 were rejected in the Office Action. Reconsideration and allowance of all claims are respectfully requested in view of the following remarks.

On page 2 of the Office Action the Examiner rejected claims 1-32 and 62 under 35 USC §101. The Examiner asserted that the claimed invention is inoperative and lacks utility. The Examiner indicated that the claims set forth a method of determining a relationship between a first textual input and a second textual input, and not a concrete method or a computer program. The Examiner further indicated that there must be a physical transformation outside the computer. The applicant respectfully disagrees with the Examiner's assertion.

The present invention relates to a computer implemented process. Therefore, the claims must be examined under M.P.E.P. §2106 for patentable subject matter for the computer related

inventions, and in particular item IV of the examination guidelines. To be statutory, a claimed computer-related process must either: (A) result in a physical transformation outside the computer for which a practical application in the technological arts is either disclosed in the specification or would have been known to a skilled artisan, or (B) be limited to a practical application within the technological arts. See Diamond v. Diehr, 450 U.S. at 183-84, 209 USPQ 1, at 6 (1981) (quoting Cochrane v. Deener, 94 U.S. 780, 787-88 (1877)) ("A [statutory] process is a mode of treatment of certain materials to produce a given result. It is an act, or a series of acts, performed upon the subject-matter to be transformed and reduced to a different state or thing.... The process requires that certain things should be done with certain substances, and in a certain order; but the tools to be used in doing this may be of secondary consequence."). It is clear from Diamond v. Diehr that the claimed invention need not produce a physical transformation outside the computer to be patentable subject matter. This physical transformation can occur inside the computer, and still remain patentable. Furthermore, a claim is limited to a practical application when the method, as claimed, produces a concrete, tangible and useful result; i.e., the method recites a step or act of producing something that is concrete, tangible and useful. See AT&T Corp. v. Excel Communications, 172 F.3d at 1358, 50 USPQ2d 1447, 1452 (Fed. Cir. 1999).

While, there is always some form of physical transformation within a computer because the computer acts on signals and transforms those signals during its operation, and hence changes the state of the computer's components during the execution of any process. Even when that physical transformation occurs within the computer, such activity is not determinative of whether the process is statutory because such transformation alone does not distinguish a statutory computer process from a nonstatutory

computer process. What is determinative is not how the computer performs the process, but what the computer does to achieve a practical application. See Arrhythmia Research Tech. v. Corazonix Corp., 958 F.2d 1053, 1057, 22 USPQ2d 1033, 1036 (Fed. Cir. 1992).

The present invention, as claimed in claims 1-32 and 62, meets the requirements of 35 USC §101 as explained in the examination guidelines as it is limited to a practical application within the technological arts. For example, claim 1 is directed to a method of determining the relationship between a first and a second textual input. Specifically, claim 1 is directed to a method for determining a relationship (such as similarity in meaning) between two textual inputs. This method assists in improving information retrieval-type tasks by identifying constituents in documents that are to be searched.

For example, the first textual input can be a search query, in natural language, and the second textural input can be the document that is searched. Claim 1 further requires the obtaining of a set of relations. This set of relations determines the way the method determines the relationship between the two documents, such as with grammatical relations or cases. See page 8 line 24. Next claim 1 requires identifying constituents in the first textual input having the obtained relations. Then claim 1 requires determining the relationship between first and second textual inputs based on the constituents identified. This relationship is the product of the claimed process. Hence the method of claim 1 can determine a relationship between the first textual input (which is the query) and a second textual input (which is one or more of the documents under consideration or indexed records of those documents) and identify those documents which have a close relationship (such as those documents which are similar in meaning) to the query, and provides the identification of those documents, or the documents themselves,

as an output to the user. See page 12 lines 5-13. This relationship is useable by either the user or another program in understanding how related or similar in meaning the two textual inputs are to each other. Therefore, it is believed that the present invention as claimed in claim 1 produces a tangible useful result, the relationship between two textual inputs, and is therefore statutorily allowable subject matter. Claims 2-22 depend from claim 1 as well, and are also believed to contain statutorily allowable subject matter. Furthermore, independent claim 28 and dependent claims 24-32 share this similar transformation, and are believed to contain similar subject matter. Claim 62 is directed to a process similar to claim 1 by determining a useable relationship between two documents, and for reasons similar to those discussed above it is believed that claim 62 pertains to statutory subject matter. Reconsideration and withdrawal of the rejection are respectfully requested.

On page 3 of the Office Action the Examiner provisionally rejected claims 1, 23, 41 and 62 under the Judicially created doctrine of double patenting over claims 1, 29, 49 and 72 of a co-pending application No. 09/097,979. With this Response a Terminal Disclaimer is submitted. Reconsideration and withdrawal of the provisional double patenting rejection are respectfully requested.

On page 4 of the Office Action, the Examiner rejected claims 1-35, 41, 43 and 62-66 under 35 U.S.C. §103(a) as being unpatentable over Turtle, U.S. Patent No. 5,265,065 in view of Liddy et al., U.S. Patent No. 6,006,221. With regards to claim 1, the Examiner indicated that Turtle taught the invention of claim 1 in its entirety with the exception of a second textual input having relations. The Examiner further indicated that Liddy et al. teaches a second textual input having relations. The Examiner also indicated that it would have been obvious to one of ordinary skill in the art to modify the communication service system of

Turtle to include Liddy's teachings by providing the second textual input that have relations. The applicant respectfully disagrees.

Turtle is directed to a system and a method for parsing a natural language textual input, and converting the input into a formatted search query. Specifically, to convert the natural language input into the formatted query the system in Turtle first takes the input phrase or query and removes all of the stop words or common words in the query. These stop words include, for example, articles such as "a" and "the", propositions such as "on", "to" and "in", and verbs such as "is", "are", "will", and "was". Once these stop words are removed, Turtle then reduces the remaining words in the textual input to their stem or root form. For example, the word "united" is reduced to its root word of "unit". Once all the remaining words in the textual input have been reduced to the respective stem or root word, the system in Turtle removes all duplicate words occurring in the query, keeping only the first occurrence of each duplicated word. For example, if the input was "The United States unites with France", the resultant textual input following parsing would be "unit state unit France." The second occurrence of the root word "unit" would be removed from the textual input of Turtle yielding a final textual input of "unit state France".

Following the stemming process, Turtle goes through each of the remaining words (which have been stemmed) in the formatted search query and pairs the first word with the next word in the chain of words in an attempt to find a phrase match in a predefined phrase database for the stemmed word pairs. If the stemmed word pairs are found in the database, the system in Turtle replaces the stemmed word pair with the phrase found in the database. Otherwise, Turtle starts over with the second word of the stemmed word pair becoming the first word in a new stemmed word pair, and repeats the process until all of the stemmed words

have become the first word in a new stemmed word pair, and repeats the process until all of the stemmed words have been analyzed for matching phrases in the phrase database. Once all of the words have been analyzed, Turtle then searches the document database for those documents that match the replaced phrases and remaining stemmed words. Turtle presents the results of that search to the user based on a relevance score. Turtle can search an entire database, or a selection of the database, such as a document summary. Further, Turtle et al. has the capability to treat non-word entries as if they were words for purposes of the searching process, for example, a non-word entry may be a legal case citation.

Liddy et al. discloses a multi-lingual document retrieval system. The system of Liddy et al. attempts to take a user query and break the user query down into language independent conceptual forms by assigning one of a predetermined plurality of group codes to each word in the input. Liddy discloses approximately 10,000 concept groups. Each concept group is purportedly a collection of words or phrases, in multiple languages, that are conceptually synonymous or nearly synonymous with each other. Usually all members of a given concept group belong to the same part of speech. Individual words that have multiple senses may occupy multiple different concept groups. As each group is considered an English language independent concept, the concept groups are not linked by broader or narrower relations.

In order to determine the concept group for a particular word in a textual input, Liddy attempts to determine the most appropriate sense for the words in the textual input. When disambiguation is needed, i.e. an input word in the sentence could belong to more than one concept group, Liddy selects the appropriate concept group using one of three sources of linguistic evidence. These sources of evidence are 1) "local

context", 2) "domain knowledge" and 3) "global information".

"Local context" is used to determine the anchor words for a given textual input. If a word within the sentence is tagged with only one concept group code, this word becomes an anchor word (unique). If more than one word in the sentence has an identical concept group code as another word in the sentence, this concept group code is also used as an anchor for the textual input (frequent). If a word in the sentence has multiple concept group codes and these group codes share a concept group code with either the unique or frequent concept group code, then that unique or frequent concept group code is selected for that particular word. However, if the word does not have an overlap with the concept group codes or any unique or frequent concept group code, the word must be evaluated according to "domain knowledge".

"Domain knowledge" uses representations to reflect the extent to which words in one concept group tend to co-occur with words of other concept groups. A sentence with a word having multiple concept categories is disambiguated to a single concept category that is most highly correlated with either a unique or frequent concept category. If several frequent anchor words exist in the textual input the ambiguous word is disambiguated to the concept category of the anchor word with the highest overall correlation coefficient. In other words, using domain knowledge, Liddy et al. attempts to select as the concept group code for an ambiguous word the code having the highest probability of being the correct concept code.

If however, during the "local context" and "domain knowledge" stages no anchor words one found, the system in Liddy proceeds to use "global knowledge". "Global knowledge" simulates the observations made in human sense disambiguation, and is based upon the assumption that more frequently used senses of words are most likely the correct sense of the word. Therefore, any words

that were not disambiguated by the "local context" or "domain knowledge" methods will have their multiple concept group codes compared to a knowledge database to determine which concept code is the most frequently used. Liddy then selects the most frequent concept code as the concept code for that word.

Once the concept codes are known for the textual input in Liddy et al., the system then searches documents in multiple databases and multiple languages for documents exhibiting concept codes that match the concept codes for the input query. Those documents matching the concept codes are then presented to the user irrespective of the language in which they were originally written. The user then may select those documents that appear most relevant to the input query, translating them as necessary.

If one were to combine Turtle with Liddy et al., as suggested by the Office Action, it would still not yield the present invention as claimed in claim 1. Specifically, the combination would take a textual input query and parse the query to remove all of the stop words and repeated words from the textual input. Following the removal of the stop words, the combined system would then proceed to stem all of the words, and remove from the input all of the duplicate words. Then the system would take the remaining words and determine the conceptual sense those words were used in. Following the system determining the concept codes, the remaining textual input for the combined system of Turtle and Liddy et al. would search a document database of multiple languages to find documents that matched the concept group codes of the first textual input query. This combined system still does not yield the system as claimed in the present invention.

For example, assuming the following hypothetical textual input in natural language: "Was there a fly on the airplane?" It is applicant's understanding that the combined system of Turtle and Liddy et al. would break the above hypothetical textual input

down to the following words "there" "fly" and "airplane" by removing the stop words. As all of the words that were identified by the combined system are in the root form, no stemming would necessarily take place.

Turning now to determining the concept group codes for the identified sentence, the combined system would likely identify under the local context a concept code anchor for the word "airplane". This is because the word "airplane" only has one sense or common usage, that being a machine that flies. The word "fly" can have at least two conceptual or sense meanings. The first meaning or sense is "to fly" as into move through the air. The second meaning or sense is an insect or an animal. As the word "fly" has at least two conceptual nodes, the combined system of Turtle and Liddy would proceed to domain knowledge in order to determine the concept group code for the word "fly". As the word "airplane" created an anchor under the local context this anchor would influence the concept code group's selection for the word "fly". As the words "fly" and "airplane" are conceptually related words under domain knowledge the combined system would select as the concept code group the group code associated with "flying" because of the influence of the anchor code for the word "airplane". Thus, under the above example the combined system of Liddy and Turtle would look for documents having conceptual codes about flying airplanes or airplanes flying. Thus, this combined system would provide back to the user results that were immaterial to the initial input query and exclude those that are material.

In contrast to the above combined system the system of the present invention would identify relevant documents to the hypothetical input query of "Was there a fly on the airplane?". The system of claim 1 starts with a "set of relations". The system of claim 1 identifies the constituents in the input query that have the identified relations. For example, assume that the

set of relations to be used in matching includes the concepts of subject and object. Then using the grammatical relation the method of claim 1 would identify the word "fly" as a subject of the sentence or the input query, and the word "airplane" as the object of the input query. Thus, the method of claim 1 would identify a relationship between the input query and documents in the database based on the word "fly" as the subject word and the word "airplane" as an object. Then the system of claim 1 searches the database and identifies in the second textual inputs those documents having the same relationships as the first textual input. The system of claim 1 returns to the user those documents pertaining to "flies on airplanes" and does not return documents relating to flying airplanes or airplanes flying.

Therefore, it is clear that the combination of Turtle and Liddy et al. would create a system that searches a database of documents in multiple languages to attempt to find documents that are conceptually related to an input query. However, the limitations of Liddy prevent the combined system from accurately identifying the correct concept of the textual input when the textual input could have multiple meanings that make "sense". However, these limitations are not present in the system claimed in claim 1 of the present invention. Therefore, it is believed that the features of the present invention as claimed in claim 1 are not taught or suggested by either Turtle or Liddy et al., taken singularly or in combination. Reconsideration and withdrawal of the rejection are respectfully requested.

Turning now to dependent claim 2 of the present invention, claim 2 adds the additional limitation that the first and second documents are compared using both the constituents and the relations. The Office Action indicated that Turtle taught this additional feature in column 13, lines 51-57. The Applicant has thoroughly reviewed the teachings of the Turtle reference, and cannot find any teaching where the system compares the two

documents using either grammatic relations, or case relations of the words. In the section cited by the Office Action the search engine relies upon only the phrase that was matched in the phrase database. However, these additional phrases are created using the same stemmed words that are used and were discussed above with regards to independent claim 1. As discussed above, Turtle does not consider the relationship between the various words or phrases in the formatted search query. Specifically, when using the phrase as suggested by Turtle the system still does not consider either the grammatical relation between the words of the phrase or the case relation between words and the input query. Liddy et al. also does not consider the grammatical relations between words in the text. As the limitations of claim 2 are not taught or suggested by Turtle or Liddy et al. it is believed that dependent claim 2 is allowable over both Turtle and Liddy et al. either independently or by virtue of its dependency from allowable claim 1. Reconsideration and withdrawal of the rejection are respectfully requested.

Turning now to dependent claim 3 of the present invention, claim 3 adds the additional limitation to claim 1 of obtaining a hierarchy of grammatical relations, and obtaining a hierarchy threshold based on a usefulness of the grammatical relations. The Office Action indicated that Turtle shows a hierarchy of grammatical relations in element 40 of FIG. 4, and obtains a hierarchy threshold based on a usefulness of grammatical relations in the hierarchy in determining the relationship between the first and second textual inputs is shown by element 44 of FIG. 4. However, Turtle does not discuss using grammatical relations, much less a hierarchy of grammatical relations. Even further from the teachings of Turtle is a threshold on such a hierarchy.

Element 40 is simply a list of stemmed words with stop words removed. There is no grammatical relation shown, much less a

hierarchy of relations. Element 44 of Turtle does not show a hierarchy threshold based on the usefulness of grammatical relations. It shows the constituent words still in stemmed form, where some of the stemmed words are grouped as phrases. The words presented in element 44 of Turtle are in the same order as the stemmed words were presented in the original query. Turtle does not consider the "relations" of these words and phrases in the formatted search query.

Furthermore, each word presented in element 44 has the same relative importance when the documents are searched for matches. Therefore, Turtle cannot determine the relative importance of each word in the list and certainly cannot do this based on the word's grammatical relation.

As neither Turtle nor Liddy et al. teach or suggest the specific limitations of claim 3 it is believed that dependent claim 3 is allowable over Turtle and Liddy et al. taken either singularly or in combination.

Furthermore, dependent claims 4-12 go on to further define the stems discussed above and are believed to be independently patentable as well. Reconsideration and withdrawal of the rejection of claims 1-12 are respectfully requested.

Claim 13 adds the additional limitation to claim 1 of obtaining a hierarchy of case information, and obtaining a hierarchy threshold based on the usefulness of a constituent having that case. Case information is defined at page 6-7 in the Specification as for example, part of speech, morphological inflections, and word boundaries. The Office Action indicated that this limitation was taught in column 15, lines 1-17 of Turtle. However, the section referred to by the Office Action shows returning results based on a partial phrase match, and not a hierarchy of case information. Nowhere in Turtle is it shown that anything is identified or decided based on a hierarchy of case information. It is simply not shown in the references. As

neither Turtle nor Liddy et al. considers the case information or the hierarchy of case information, it is believed that dependent claim 13 is allowable over Turtle and Liddy et al. Furthermore, dependent claims 14-22 are believed to be independently allowable over Turtle and Liddy et al. because they further define the previous claims and are neither taught no suggested by the references. Reconsideration and withdrawal of the rejection of claims 13-22 are respectfully requested.

Claim 23 recites a method for determining a relationship between first and second textual inputs including "analyzing the first textual input to obtain relations of constituents thereof." Further independent claim 23 requires determining a relative importance of the constituents in determining the relationship between the first and second textual inputs based on the relations obtained, and determining the relationship between the first and second textual inputs based on the constituents and the relative importance of the constituents.

The Office Action indicated that the limitations of independent claim 23 were taught by the Turtle reference in combination with the Liddy et al. reference. However, as discussed above with regard to independent claim 1 and dependent claims 3 and 13, neither Turtle nor Liddy et al. rank constituents in any manner, and in particular do not rank the constituents based on their relations or relative importance. Furthermore, neither Turtle nor Liddy et al. determines a relationship between the first and second textual inputs based on the relative importance of the constituents. In fact all Turtle does is present to the user the results of a comparison of documents based on the number of words in the search query that appear in the document as either individual words or phrases. All Liddy et al. does is present to the user documents in multiple languages that match the determined concept code. As neither Turtle no Liddy et al. teach the limitations of claim 23, it is

believed that independent claim 23 is allowable over Turtle. Furthermore, dependent claims 24-32 are believed to be independently allowable independent claim 23 in ways neither taught nor suggested by the references. Reconsideration and withdrawal of the rejection of claims 24-32 are respectfully requested.

Claim 33 requires that the plurality of constituents have a predetermined usefulness in determining the relationship based on relations of constituents in the textual material. As discussed above, neither Turtle nor Liddy et al. teach or suggest determining the usefulness of the constituents in determining the relationship between two textual inputs. As neither Turtle nor Liddy et al. discloses a computer readable medium having a plurality of constituents having a predetermined usefulness, it is believed that independent claim 33 is allowable over Turtle and Liddy et al. Furthermore, dependent claims 34 and 35 are believed to be independently allowable. Claim 34 is drawn to the index including syntactic structure indicative of a constituents' grammatical relation, and claim 35 states that the usefulness is based on a portion of the grammatical relation along a hierarchy. Of course, neither of these are taught of suggested by the references. Reconsideration and withdrawal of the rejections of claims 33-35 are respectfully requested.

Claim 41 is directed to a computer readable medium storing a data structure used in determining a relationship between the first and second textual inputs. The data structure comprises a plurality of pre-computed aspects of at least one of the first and second textual inputs. The pre-computed aspects are useful in determining a relationship between the first or second textual inputs. Further claim 41 includes the additional limitation of that plurality of pre-computed aspects includes a linguistic analysis of at least a portion of the first and second textual input. Turtle does not teach or suggest use of precomputed

linguistic analysis as a factor in a data structure used in determining the relationship between first and second textual inputs. Therefore, it is believed that independent claim 41, is allowable over both Turtle and Liddy et al.

Furthermore, claim 43 includes within the data structure a plurality of constituents wherein those plurality of constituents have a predetermined indication of usefulness. As discussed above, it is believed that Turtle does not teach or suggest these limitations of claim 43. Therefore it is believed that dependent claim 43 is independently allowable and is further allowable by virtue of its dependency from allowable independent claim 41. Reconsideration and withdrawal of the rejection are respectfully requested.

Claim 62 is directed to a method for determining a relationship between the first and second textual inputs. The method comprises obtaining a hierarchy of relations, obtaining a hierarchy threshold based on the usefulness of the relations. The method of claim 62 further comprises identifying constituents in the first textual input that have the relations in the hierarchy, and determining the usefulness of the identified constituents by locating the identified constituents in the hierarchy. Finally, the method of claim 62 determines the relationship between the first and second textual inputs based on the identified constituents having an associated relation above the hierarchy threshold. As discussed above with regards to claim 1 neither Turtle nor Liddy et al. teach or suggest a hierarchy of relations. Furthermore, neither Turtle nor Liddy et al. teach obtaining a hierarchy threshold based upon a usefulness of the relations in determining the relationship between the first and second textual inputs. Therefore, it is believed that independent claim 62 is allowable over both Turtle and Liddy et al. Reconsideration and withdrawal of the rejection are respectfully requested.

Turning now to dependent claims 63-66, claims 63-66 are dependent from claim 1, and contain the additional limitations of obtaining case information based on the usefulness of a constituent having that case, and obtaining grammatical relation information based on the usefulness of a constituent having that grammatical relation. Claims 63-66 have the limitations of both dependent claims 3 and 13 for determining the relationship between a first and second textual input. As the limitations contained in both dependent claims 3 and 13 are recited in dependent claim 63, and dependent claims 3 and 13 are believed independently allowable over Turtle and Liddy et al.; the combination of the limitations of claims 3 and 13 is allowable as well over Turtle and Liddy et al. Furthermore, it is believed that dependent claims 64-66 are independently allowable as well because they are neither taught nor suggested by the references.

Reconsideration and withdrawal of the rejection are respectfully requested.

In conclusion it is believed that claims 1-33, 41, 43 and 62-66 are allowable over Turtle and Liddy et al. Reconsideration and allowance of claims 1-35, 41, 43 and 62-66 are respectfully requested.

The Director is authorized to charge any fee deficiency required by this paper or credit any overpayment to Deposit Account No. 23-1123.

Respectfully submitted,

WESTMAN, CHAMPLIN & KELLY, P.A.

By: _____

Nathan M. Rau, Reg. No. 45,466
Suite 1600 - International Centre
900 Second Avenue South
Minneapolis, Minnesota 55402-3319
Phone: (612) 334-3222 Fax: (612) 334-3312

NMR/jme